JVI Journals.ASM.org

GENOME ANNOUNCEMENT

Complete Genome Sequence of Human Respiratory Syncytial Virus Genotype A with a 72-Nucleotide Duplication in the Attachment Protein G Gene

Wan-Ji Lee,[a] You-jin Kim,[a] Dae-Won Kim,[b] Han Saem Lee,[a] Ho Yeon Lee,[a] and Kisoon Kim[a]

Division of Respiratory Viruses, Center for Infectious Disease,[a] and Systems Biology Team, Center for Immunity and Pathology,[b] Korea National Institute of Health, Cheongwon-gun, Chungbuk-do, Republic of Korea

The complete genome sequence of human respiratory syncytial virus genotype A (HRSV-A) with a 72-nucleotide duplication in the C-terminal part of the attachment protein G gene was determined and analyzed. The genome was 15,277 bp in length, and 0.46 to 6.03% variations were identified at the nucleotide level compared with the previously reported complete genome of HRSV-A. Characterization of the genome will improve understanding of the diversity of the HRSV-A major antigens and enable an in-depth analysis of its genetics.

Human respiratory syncytial virus (HRSV), known to be a leading cause of severe respiratory infections in neonates and children, belongs to the *Pneumovirus* genus within the family *Paramyxoviridae* (1). HRSVs have enveloped, negative-sense, single-stranded RNA genomes of approximately 15 kb that encode 11 viral proteins. The HRSVs are divided into two groups, A and B, based on variability in antigen reactions against the attachment (G) and fusion (F) glycoproteins (4). To date, 10 HRSV-A genotypes have been designated GA1 to GA7, SAA1, NA1, and NA2. The HRSV-B genotypes include GB1 to GB4, SAB1 to SAB3, and BA1 to BA6 (7). Since 1999, a duplication of 60 nucleotides (nt) in the C-terminal third of the G gene from HRSV-B has been identified in South America (Argentina), and this has now spread worldwide (6). Interestingly, a similar but unique duplication was reported in 2012 from partial sequence information about the G gene in HRSV-A, which is characterized by a 72-nt duplication in the C-terminal part of the G gene (2).

In this study, complete sequence information on a putative newly emerged HRSV-A strain (HRSV-A/GN435/11) with a 72-nt duplication in the C-terminal of the G gene has been determined. It was isolated from elderly patients with acute respiratory illness in Korea. Full-length viral RNAs were prepared from supernatants of Hep2 cell inocula, and the first-strand cDNA was synthesized by random priming. Entire sequences were obtained from subclones of PCR fragments corresponding to dozens of inner primers by using an ABI 3730xl DNA sequencer (Applied Biosystems, Foster City, CA). Sequence assembly and consensus calling were performed to produce complete genome information using SeqMan Pro of the Lasergene 8 program suite (version 8.0; DNAStar, Madison, WI).

The entirety of this newly identified HRSV-A/GN435/11 strain was 15,277 bp, including the 3′ leader and 5′ trailer sequences. From this, characteristic open reading frames (ORFs) of 11 encoded viral proteins were deduced for the NS1, NS2, N, M, P, G, F, SH, M2-1, M2-2, and L genes. The nucleotide sequence variability of HRSV-A/GN435/11 ranged from 0.46% to 6.03% compared with previously deposited whole-genome information for HRSV-A (3, 5). The deduced amino acid sequence identity ranged from 84.1% to 100%, in accordance with the ORFs. A nonsynonymous start codon variation (ATG to ACG) was identified, with alternative start codons six bases downstream in the M2-2 gene. Synonymous stop codon variations (TAA to TGA) were found in the NS2 and F genes. The smallest G-to-C ratio was observed in the M2-2 gene (28.84%), and the overall G+C content was 33.44%. Phylogenetic analysis showed that the G gene, with a 72-nt duplication in HRSV-A/GN435/11, was significantly different from previously reported HRSV-A strains and clustered together with the HRSV-A/ON1 strain (2).

The fundamental genomic information presented in this study will help clarify the antigenic variability and dynamics of HRSV evolution by enabling in-depth analyses of changes in the entire HRSV genome sequence.

Nucleotide sequence accession number. The complete sequence of the novel HRSV-A/GN435/11 strain was submitted to GenBank and assigned accession no. JX627336.

ACKNOWLEDGMENT

This study was supported by the Intramural Research Fund (2010-N47001-00) of the National Institute of Health of Korea.

REFERENCES

1. **Anderson LJ, Parker RA, Strikas RL.** 1990. Association between respiratory syncytial virus outbreaks and lower respiratory tract deaths of infants and young children. J. Infect. Dis. **161**:640–646.
2. **Eshaghi A, et al.** 2012. Genetic variability of human respiratory syncytial virus a strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. PLoS One **7**:e32807. doi:10.1371/journal.pone.0032807.
3. **Kumaria R, Iyer LR, Hibberd ML, Simoes EA, Sugrue RJ.** 2011. Whole genome characterization of non-tissue culture adapted HRSV strains in severely infected children. Virol. J. **8**:372. doi:10.1186/1743-422X-8-372.
4. **Mufson MA, Orvell C, Rafnar B, Norrby E.** 1985. Two distinct sub-

Received 19 September 2012 Accepted 19 September 2012
Address correspondence to Kisoon Kim, tigerkis@gmail.com.

doi:10.1128/JVI.02571-12

types of human respiratory syncytial virus. J. Gen. Virol. **66**(Pt 10): 2111–2124.
5. **Rebuffo-Scheer C, et al.** 2011. Whole genome sequencing and evolutionary analysis of human respiratory syncytial virus A and B from Milwaukee, WI 1998-2010. PLoS One **6**:e25468. doi:10.1371/journal.pone.0025468.
6. **Trento A, et al.** 2010. Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. J. Virol. **84**:7500–7512.
7. **Venter M, Madhi SA, Tiemessen CT, Schoub BD.** 2001. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. J. Gen. Virol. **82**:2117–2124.